Edward Hribal

Professor Fukuyama

STAT-S 470

April 9, 2021

**Mini Project 2: Political Polarization**

**Question Set 1:**

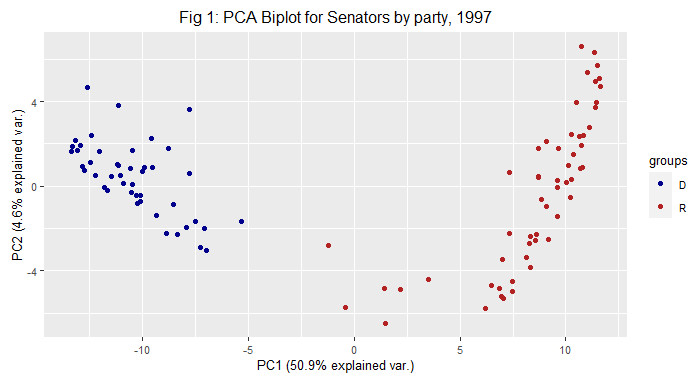
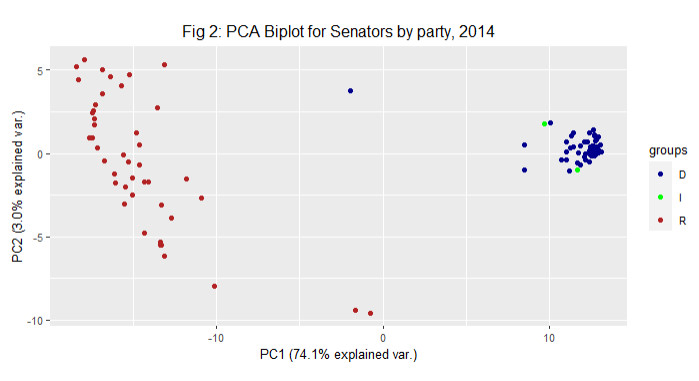
Polarization in two years: Take a look at the voting records for one of the earlier sessions in the dataset and one of the later sessions (e.g., 1989 and 2014). Using either PCA on the senators by bills matrix or multi-dimensional scaling on a measure of distance between senators that you construct, make a plot describing the relationships between the senators. Does it look like they fall on a one-dimensional liberal/conservative axis? Are there outliers? Are the patterns similar in the two years that you chose? What are the differences in the PCA/MDS plots in the two years that you chose?

To observe the voting patterns of senators from an earlier and a later session of the datasets provided, I first chose the 1997 and 2014 voting records. The 17-year period between the votes is not the widest range I could have chosen from the dataset; however, I chose the two periods because of the similar composition of congress during each year. Each had Republican controlled congresses in the middle of a Democratic president’s second term, and along with avoiding the noise of votes in 1998 regarding Clinton’s impeachment, the similar congressional properties in each year should isolate the effect of polarization as much as possible.

In order to measure the distance between senators in each of the years chosen, I used a PCA biplot to demonstrate the distance between each senator’s voting patterns, and grouped by party affiliation. I chose to plot this with a PCA as opposed to an MDS because I wanted to preserve the actual magnitude of the distances between the senators’ votes. Whereas MDS would have greatly reduced the visibility of outliers, the actual distance of a senator’s voting record is important to preserve in this visualization, as the outliers are likely to be important in this particular data analysis.

When constructing the PCA plot for 1997, the scree plot showed that a significant amount of the variance in the votes is explained by the first principal component. This indicates that a one-dimensional structure could sufficiently demonstrate the variance in the data, however, there is enough variance in the second x value of the scree plot that the elbow test could justify a two-dimensional structure. For the 2014 data, the scree plot shows overwhelming variance in the first principal component, with the second component very close to zero alongside each x value following. This is already a sign that polarization has grown significantly between 1997 and 2014. Since the second principal component had a fair amount of variance in the 1997 data, it indicates that there are some additional principles that may guide the voting patterns of senators beyond a leftward/rightward bent towards their respective parties. However, since a very significant portion of variance can be explained with only the first principal component in 2014, it indicates that there is very little else that can describe a senator’s voting patterns besides what party they are in.

To further explore this, I created a biplot for both 1997 and 2014 without any variable axes, each grouped by the party each senator belongs to. The Democratic senators are marked by blue, and Republican senators are marked by red. Also, after 1997 two independent senators were elected and held office in 2014. They are marked by green in the biplots shown below.

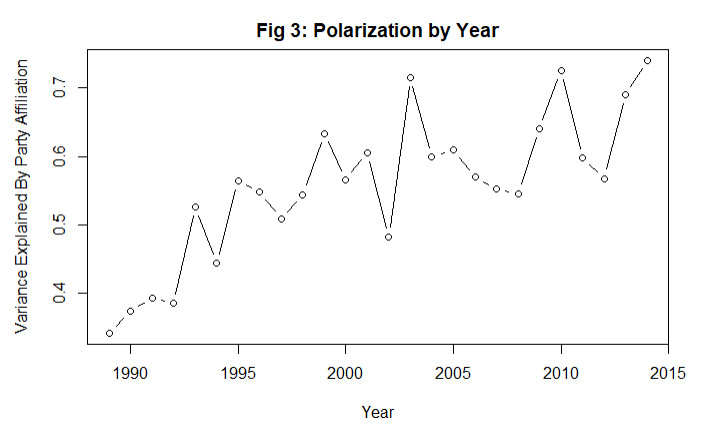


As noted on the figure’s principal component axes, the variance is explained by a single dimension to a much greater degree in 2014 than in 1997, with 74.1% of the variance being explained by a one-dimensional difference in 2014 compared to 50.9% in 1997. This trend can also be observed in the clustering of the datapoints in the 2014 biplot, as opposed to the more spread-out data in 1997. In the earlier dataset, the datapoints for each party is much more vertically spread and more angular than the more recent data. Also, there were more notable centered outliers in 1997 than in 2014, as there were about 7 outliers in the center of the dataset in 1997, and only 3 in 2014.

**Question Set 2:**

Polarization over time: Next, we want to look at this more systematically over time. Repeat the same analysis you used in the first part for every year in our dataset. Construct a measure describing polarization, and plot it over time (think about variance explained by the first principal component, average distance between the scores for Democrats and Republicans, or similar measures). If you are able to make a clean visualization showing the PCAs or MDSs for each session of congress and you think it helps to support your claims, include that as well.

Do you think that polarization has increased over time? Is something more complicated happening? Back up your claim with specific aspects of your analyses and graphs, and describe why they support your claim.

In repeating the steps from the first question set onto every year included in the dataset, it is clear that polarization has increased in congress significantly over time. Using the variance explained by the first principal component as a benchmark for polarization in a given year, I plotted each of the values to each year they correspond to in a lined scatterplot below.

Based on an understanding of events in American politics the last 25 years, it seems like using the PC1 variance as a measure of polarization can present an accurate reflection of the divide in congress over time. There is first a spike around 1994, when Republicans began promoting their “Contract with America” agenda and sought uniformity within the party, then another spike in the late 90’s around the time of Clinton’s impeachment. This was followed by a significant drop in 2002, as lawmakers briefly united for a series of bills following the September 11th attacks, before the Iraq War drew congress back into polarization. While there was a surprising series of declines in the partisan-explained variance in the mid-2000’s, it quickly rose again after the election of Barack Obama in 2008 and the obstructionism that came afterwards.

Overall, it is clear that polarization has risen in the last 25 years intuitively and based on the PC1 variance metric, and occurrences within the chart that cause it to deviate from a linear path have valid explanations within the context of the time period.

**Question Set 3:**

Ideological Position of One Senator: Choose one senator who served in all (or at least many) of the years we have represented in our dataset, and make a plot describing his or her position relative to the two major parties over time. Describe the implications of your plot either in terms of which party is more to blame for polarization or relating to the question of whether that senator’s politics has changed over time. You should also describe the limitations of your analysis, and whether there are other possible explanations for the phenomena you see.

I selected Mitch McConnell to be the single senator I analyzed between 1989 and 2014, as his leadership positions within the Republican party make him an interesting figure to observe with his proclivity to vote with the party as a whole, as well as his distance from Democrats, over time. To do this, I calculated the average PC1 and PC2 value for each party in each year, then found the Euclidian distance between McConnell’s point on the ggbiplot and the average value in each year. A larger distance equates to a year in which he differed more with the respective party’s consensus, and a smaller distance means he voted more in line with the party. His distance over time from the GOP and Democratic average is plotted below.

Chart, line chart

Description automatically generated

Chart, line chart

Description automatically generated

While there is significant inter-year variation, particularly between McConnell and the Democrats, it is clear that over time he has come to align himself more with Republicans, and less with the Democratic party. There is significant inter-year variation, which may indicate years with particularly heightened tensions between the two parties, but it appears that McConnell’s politics have changed over time to reflect more of the GOP’s consensus and less of the Democratic party’s voting patterns. Since McConnell is a longtime Republican and currently the party’s leader, this would seem to indicate that the Republican party is at fault for the increasing polarization over the last 25 years.

However, there are limitations to this analysis which may cast doubt on broad conclusions such as the one above. For example, the GOP and McConnell are clearly farther away from the Democratic party in 2014 than 1989, but it can be difficult to assign blame because the substance of the bills being voted on is not included as a variable. It is difficult to tell if Republicans are no longer voting for centrist policies they previously voted for or vice versa, the biplots only give us a relative gauge of how frequently the Democrats and Republicans stick together without consideration for what they are sticking together on.

Taking an average value of each party’s votes from the principal components may not give us an accurate reflection of which party is causing the most polarization either. Using Figure 2 as an example, it would appear that the Democrats are contributing heavily to polarization, as their observations are all clumped together far away from the center with few outliers. However, while the GOP has more variance in their members towards the center, the party also has many more points farther from the center compared to the Democrats. Putting the data into the context of the real world in which the most controversial voices are often the loudest and most influential, it should be reasonable to put more weight on more extreme principal component values in calculating a party’s mean based on the impact they have.

Another limitation with this kind of analysis is that the distances between the GOP’s PCA components and McConnell’s do not reflect which of the two did the most “moving” in their voting patterns. It is possible that the party simply became more like McConnell rather than the other way around. The timing of polarization and McConnell’s proximity to the GOP norm is also an interesting point to note though, as there are years such as 1993 and 2009 in which the polarization in Figure 3 increased, while in Figure 4 McConnell’s distance from the party average increased. This would seem to indicate a lag in which the Republican party went farther to the right, and McConnell did not ‘catch up’ until a year or two later.